增强现实技术赋能移动新闻客户端初探索

李　磊
（天目传媒有限公司，浙江　杭州　310039）

摘　要：2020 年 9 月，中共中央办公厅、国务院办公厅印发了《关于加快推进媒体深度融合发展的意见》中明确要求以先进技术引领驱动融合发展。本文探索如何利用增强现实技术为移动新闻客户端赋能，从而实现新闻内容的精细化、互动化和沉浸化，助力媒体深度融合。

关键词：增强现实（Augmented Reality）；虚拟场景；沉浸式体验；媒体深度融合　**中图分类号：**G206　**文献标识码：**A
文章编号：1671-0134（2021）09-155-04　DOI：10.19483/j.cnki.11-4653/n.2021.09.051
本文著录格式：李磊 . 增强现实技术赋能移动新闻客户端初探索 [J]. 中国传媒科技，2021（09）：155-157，42.

增强现实技术（Augmented Reality，简称 AR）是虚拟场景与现实场景的融合技术，[1]20 世纪五六十年代，美国摄影师兼发明家 Morton Heilig 用他拍摄电影的经验，发明了一种名叫“Sensorama Stimulator”的机器，它可使用图像、声音和震动让使用者感受骑着摩托车风驰电掣的场景。以此为契机，AR 也拉开了它的发展史，AR 技术发展至今具备虚拟现实融合、实时交互、环境感知等特征，尤其是三维跟踪、三维注册是其最重要的特征。目前多用于游戏、文旅、医疗、工业、军事等应用场景。

几年前一款风靡全球的手机游戏《Pok é mon GO》让人们第一次在个人移动终端上感受到了 AR 的奇妙之处，让普通用户了解了增强现实这个概念。这里所理解的增强现实，是用虚拟内容来做视觉上的增强，通过手机屏幕或专业眼镜设备（如：Microsoft HoloLens、Magic Leap、Google Glass）来重现。AR 技术的关键在于设备对周围环境的感知理解，最基础的是确定设备自身的空间位置，再进阶的是对环境进行实时重建（SLAM）；目前最高级的是识别、认知和交互。定位是基础，只有设备自身的位置精确地被定位，虚拟内容才能和现实场景很好地、实时地进行结合，以达到足够的真实感。AR 主要是对现实场景的增强，而非完全的虚拟化，让虚拟的事物融入现实场景且定位在三维空间里，虚拟物体不会随着观看者的移动而移动。

1.AR 技术原理

目前，AR 需要通过外部设备来实现，如手机，通过获取摄像头拍摄现实场景图像并在上面叠加虚拟物体，或通过专业的 AR 设备，如 AR 眼镜、AR 头盔，也可将虚拟物体成像在眼镜里。[2]AR 从其技术实现原理上可以分为两大类。

1.1　利用计算机视觉技术来实现 AR

利用计算机视觉来实现 AR，主要是利用 CV 算法把现实世界场景与屏幕（如手机屏幕、AR 眼镜）之间形成映射关系，使得屏幕中重现的现实场景上叠加虚拟三维场景，从技术实现方式上可分为两类。

第一种是 Marker-Based AR，此方法必须先预设一个基础的标记物，如一张名片或一张海报，然后把该标记物放到现实场景中的任何位置，这等于在现实场景中定位了一个平面，然后通过手机或 AR 眼镜的摄像头对标记物进行识别，CV 算法会将标记物中心为原点的坐标系作为模板坐标系，从而将模板坐标系和屏幕坐标系建立映射关系，根据这个映射关系，在屏幕上显示出的虚拟三维场景可以实现依附在标记物上的效果。

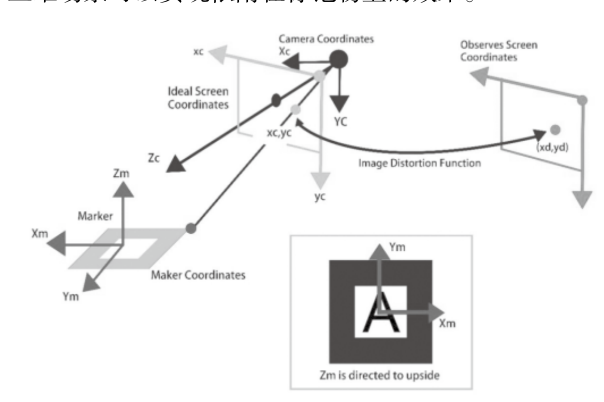


图 1 Marker-Based AR 实现原理

第二种 Marker-Less AR，其实现原理与上述第一种相比，不需要预先制作特殊标记物，只需具备任何有足够多特征点的物体。比如：身体、人脸、手势、车辆等各类物体作为基准，通过机器学习（如 SURF、ORB、FERN 等算法）来提取物体特征点，并且记住这些特征点。当手机或 AR 眼镜的摄像头扫描场景，会提取场景内的物体特征点并与机器学习记住的物体特征点进行逐一比对，如果扫描到的特征点和记录的特征点匹配数量达到或超过阈值，系统会判断为识别到该物体，然后建立映射关系，在屏幕上就可以将虚拟场景依附在各类物体上，美颜相机中人脸三维动效贴纸效果就是这样实现的。

1.2　基于地理位置信息来实现 AR(LBS-Based AR)

其原理是通过移动设备的 GPS、罗盘、陀螺仪来获

取用户所在位置的经纬度、方位和设备倾斜角度，然后从数据源（如：在线地图）获取该位置附近物体，如景点、学校、餐饮、车站等目标 POI 信息，通过这些信息建立目标物体在现实场景中的平面基准,之后再通过坐标系变换，建立映射关系，以此在屏幕中呈现虚拟场景与现实场景的组合、叠加。此方案利用移动设备的 GPS、罗盘、陀螺仪等传感器来实现，摆脱了对固定标记物的依赖，不用实时识别标记物和计算特征点，在移动端上系统性能和用户体验也要好于 Marker-Based AR 和 Marker-Less AR。

2.AR 主流技术解决方案

一套完整的移动端 AR 技术解决方案主要包含 AR 引擎（SDK）、AR 内容制作、AR 内容识别、AR 内容管理等四大业务模块，其中最为关键的是 AR 引擎，它是决定 AR 系统是否能实现现实空间叠加虚拟场景、SLAM（运动跟踪定位），光照估计、阴影等功能，同时又可基于人工智能技术，通过人脸识别、图像识别、动作识别等方式，智能分析认知现实场景，达到增强现实效果。

大多数 AR 引擎都是以 Google ARCore、APPLE ARKit 为基础，它们分别用于支持 Android 设备和 IOS 设备，并在基础上再次封装扩展开发，以适配更多的移动设备和应用系统，市场主流的 AR 引擎主要有以下几种。

（1）ARCore：Google 公司研发的 AR 引擎，仅可用于开发 Android 平台上的 AR 应用，不支持 IOS 设备，主要实现运动追踪、平面追踪、点云图、云锚点、光照估计、环境探针、人脸追踪、2D 图片追踪、人物遮挡等。

（2）ARKit：苹果公司研发的 AR 引擎，只支持 IOS 设备，用于开发 iPhone 和 iPad 平台上的 AR 应用，与 ARCore 一样能实现运动追踪、3D 物体追踪、平面追踪、点云图、云锚点、光照估计、环境探针、人脸追踪、动作捕捉等。[3]

（3）Vurforia：高通研发的 AR 引擎，它包含了 ARCore、ARKit，同时支持 Unity3D 创作的虚拟场景内容，支持 Android、IOS 设备，该引擎产生的业务数据可在云端存储。

（4）EasyAR：视辰信息科技研发的 AR 引擎，它包含了 ARCore、ARKit，同时也支持 Unity3D 创作的内容，根据应用场景自研了一套 SLAM，整套 AR 系统支持本地、云端部署。

（5）阿里 AR：阿里巴巴研发的 AR 引擎。主要是在 2D 图像识别，3D 物体追踪功能，该引擎常见的应用场景是支付宝过年扫“福”活动。

（6） LandMarkAR：是字节跳动研发的 AR 引擎，主要应用在抖音 App 内拍摄短视频为城市地标增添 AR 特效场景。

综上所述，主流 AR 引擎解决方案都有各自的能力特点，也存在一定的适配和应用场景限制。以移动客户端为例，完整的 AR 技术解决方案选型应考虑以下几个方面。

（1）引擎通用性和设备适用性：引擎能够同时支持 ARkit 和 ARcore 双平台，支持大部分 Android 和 IOS 设备，如有游戏需求需支持 Unity。

（2）功能的丰富性：能支持表面跟踪、平面图像跟踪、3D 物体跟踪、运动跟踪、人脸跟踪，实现光照估计、云锚点、点云图等功能。

（3）系统接入的易用性：采用 SDK 的方式接入移动应用或以 Web 的方式 实现扫描识别图和物体来呈现 3D 动画模型、视频、图片、文字、UI 按钮等效果，并且支持 3D 模型交互。

（4）内容的管理和分发：支持 AR 文件的云端存储、处理、审核，以及支持 CDN 分发。

（5）AI 的扩展性：可以通过 AI 智能分析认知现实场景及算法训练，比如表情、手势、身体姿态等，从而做对应的增强显示，达到增强现实效果。

以天目新闻客户端为例，经过充分的技术调研，结合客户端内容定位和运营需求，采用 EasyAR 的整体技术解决方案，该方案包括四个主要部分：AR 引擎（移动端 SDK）、内容脚本制作系统、云端识别管理系统（Cloud Recognition Service）、云端内容管理系统（Operation Center）。

AR 引擎（移动端 SDK）可支持 IOS 和 Android 两个平台，适配目前市面上主流机型，尤其对众多 Android 机型适配友好。该引擎主要能实现下列功能。

（1）平面图像的识别和跟踪：实时识别与跟踪有纹理的平面物体，比如一本书、一张名片，甚至是一面涂鸦墙，从标准图像动态生成跟踪目标，同时识别和跟踪多个目标。

（2）3D 目标的识别和跟踪：实时识别和跟踪自然场景中有丰富纹理的三维物体。目标物体可以是不同形状和结构，且可同时识别和跟踪多个 3D 物体。

（3）稠密地图空间定位与跟踪：利用手机相机或 AR 眼镜相机对周围环境进行三维稠密重建，得到稠密的点云地图和网格地图，利用稠密空间地图让虚拟物体更好的融入真实环境中，以实现真实物体和虚拟物体间的正确遮挡、碰撞等 AR 虚拟场景。

（4）稀疏地图空间定位与跟踪：利用手机相机或 AR 眼镜相机扫描周围环境，构建三维环境点云，每一个三维点都记录周围的局部视觉信息，最终生成环境的三维视觉地图并提供视觉定位跟踪功能，适用于开发持久化 AR 应用或多人互动 AR 应用。

（5）运动跟踪：可持续追踪移动设备在空间中的位置和姿态，将虚拟物体和真实场景实时对齐于同一坐标系，可用于 AR 展示、AR 游戏、AR 视频或拍照等应用，可以体验到虚拟场景和真实场景融合在一起的感受。如有持久化 AR 应用需求，可结合 EasyAR 稀疏空间地图定位与跟踪一起使用。

除上述主要功能外，该引擎还支持提供多种编程语言接口，尤其提供了对原生开发友好的 Java 和 Objective-C 编程接口，支持录制视频，内置视频播放器且支持 H.264、H.265 硬件解码，支持 3ds Max、Maya、Unity 等主流三维场景制作工具，针对移动端网络情况支

持多种压缩算法，提升用户体验。

内容脚本制作系统支持在线内容编辑和更新，包括模型编辑器、场景编辑器和脚本编辑器，支持通用javascript 脚本，从简单的模型展示到复杂的小游戏都可以通过脚本编辑器来完成。最大的优点是 AR 内容如果有变动，只需更新 AR 内容无需更新客户端，避免了客户端打包上架发布的繁琐过程。

云端识别管理系统（Cloud Recognition Service）会将图像信息作为识别请求发到云端识别，服务器检索到关联图库里与之匹配的目标图像，然后调用引擎加载三维虚拟模型等，渲染出相应的 AR 效果。该系统可实现平面目标识别、手势识别、姿态识别等多种 AR 触发方式，主要由识别系统、监控系统、统计系统等业务子系统构成，其中识别系统是 CRS 的核心组件，包括管理识别目标 API 的 Targeter 系统和 Searcher 系统，以及底层基于 CV 特征的检索系统 Retriever 系统。同时，还能为业务使用方提供监控、统计业务，用于监控识别扫描命中分析及活跃 AR 扫描设备情况。

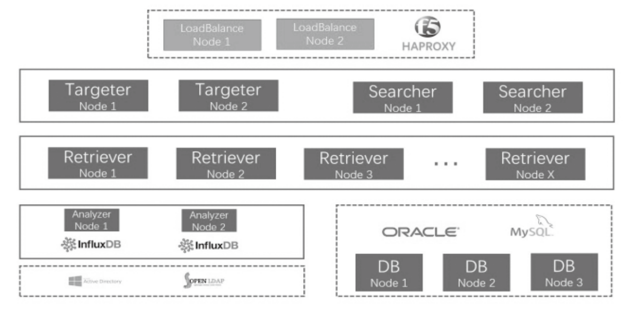


图 2　系统部署图（CRS 系统与所依赖的系统之间的关系）

云端媒资内容管理系统 (Operation Center) 是提供给 SDK 相配套的 AR 内容包进行存储管理、识别图管理、版本管理、内容分发、CDN 加速等，同时与云端识别管理系统协同工作，为有内容安全管理要求的用户提供了一套完整的云端私有化部署系统。

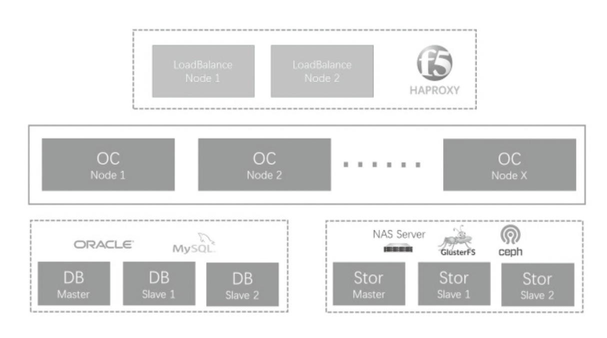


图 3　系统部署图（OC 系统与所依赖的系统之间的关系）

3.AR 技术实现内容融合创新

2020 年 9 月，中共中央办公厅、国务院办公厅印发的《关于加快推进媒体深度融合发展的意见》中明确指出要以先进技术引领驱动融合发展，用好 5G、大数据、云计算、物联网、区块链、人工智能等信息技术革命成果，加强新技术在新闻传播领域的前瞻性研究和应用，推动关键核心技术自主创新。[4] 要利用 AR、VR、大数据、5G 等技术赋能内容跨媒体制作与呈现，实现内容传播精细化与沉浸化。

2018 年 4 月，新华社客户端发布 AR 新闻《习近平的最大爱好》，用户使用 AR 功能扫描二代身份证，就可以成功进入 AR 新闻（如下图所示）。




2021年 3 月，天目新闻客户端正式上线了端内原生 AR 功能，结合当时《寻找浙江金名片》的主题报道活动，开发了 AR 内容和运营活动场景，用户只需打开客户端内“扫一扫”功能，拍摄“比心”手势，即可呈现出“浙江九张金名片”，并可参与有奖答题活动，整个主题系列内容在全网分发 200 万次以上（如下图所示）。



AR 技术给融媒体时代带来的是一种全新的、前所未有的沉浸感，它将虚拟内容与现实环境相结合，以达到扩展新闻内容信息量、增强用户交互，从而让用户直接感知并触达新闻现场，这是 AR 新闻内容最突出的特点。

AR 新闻为用户带来强烈的交互感，提升新闻的趣味性。文字调动的是用户的视觉，电视调动的是用户的听觉和视觉，AR 技术在内容触达上直接调动了用户的视觉、听觉、触觉等多种感官，在新闻传播中创新实现了多感官联动感知内容的形态，让用户沉浸于新闻现场，体验“最真实”的新闻现场感，虽然 AR 新闻部分提供的是虚拟场景，但它调动用户多感官的感知却是无比真实的。同时，内容被 AR 技术赋能后，借助强交互性的特点，以游戏的方式增加新闻内容的趣味性，使年轻用户更愿意主动去接受 AR 新闻，从而提升新闻内容的传播力。

AR 新闻内容不再是一维或二维符号的叙述，而是被三维甚至四维的方式立体地还原了新闻发生的现场。AR 新闻不是转述新闻现场以及事件的详细信息，而是直接传递新闻现场给用户，让用户自行体验现场、感受现场，对新闻事件的了解和感受远比传统新闻时期来得真实和

（下转第42页）